

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 17, 2004, 18:47:23 ; Search time 25.1148 Seconds
(without alignments)
6683.183 Million cell updates/sec

Title: US-09-989-981A-2
Perfect score: 3369
Sequence: 1 MGELPFLSPGARGPHINRG.....PALVILGIVIKVDYLISR 652

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1045404 seqs, 257433775 residues

Total number of hits satisfying chosen parameters: 1045404

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09A_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3369	100.0	652	9	US-09-837-992-1
2	3369	100.0	652	10	US-09-989-981A-2
3	2744.5	81.5	651	9	US-09-837-992-3
4	2744.5	81.5	651	10	US-09-989-981A-6
5	2744.5	81.5	651	14	US-10-090-455-6
6	1177	34.9	256	15	US-10-104-047-2795
7	701.5	20.8	672	10	US-09-989-981A-4
8	693.5	20.6	655	10	US-09-961-086-1
9	693.5	20.6	655	15	US-10-405-806-13
10	691.5	20.5	655	9	US-09-981-353-35
11	691.5	20.5	655	14	US-10-120-687-61
12	691.5	20.5	655	15	US-10-405-806-2
13	689.5	20.5	655	9	US-09-866-866A-10
14	689.5	20.5	655	14	US-10-090-455-5
15	688.5	20.4	673	10	US-09-989-981A-8

16	688.5	20.4	673	14	US-10-090-455-7	Sequence 7, Appli
17	683.5	20.3	655	9	US-09-866-866A-27	Sequence 27, Appli
18	677	20.1	657	9	US-09-866-866A-14	Sequence 14, Appli
19	628	18.6	1095	15	US-10-369-493-2025	Sequence 2025, Ap
20	619	18.4	725	12	US-10-424-599-175941	Sequence 175941,
21	602.5	17.9	1049	15	US-10-369-493-1520	Sequence 1520, Ap
22	592.5	17.6	674	14	US-10-090-455-4	Sequence 4, Appli
23	592.5	17.6	674	16	US-10-429-160-10	Sequence 10, Appli
24	586	17.4	663	13	US-10-108-605-345	Sequence 245, App
25	584.5	17.3	1084	12	US-10-424-599-342078	Sequence 242078,
26	584.5	17.3	1101	12	US-10-425-114-63125	Sequence 63125, A
27	580	17.2	638	13	US-10-072-621-10	Sequence 10, Appli
28	579.5	17.2	658	15	US-10-369-493-5347	Sequence 5347, Ap
29	576.5	17.1	646	13	US-10-072-621-9	Sequence 9, Appli
30	576.5	17.1	646	14	US-10-090-455-2	Sequence 2, Appli
31	574.5	17.1	599	15	US-10-210-130-14	Sequence 14, Appli
32	573.5	17.0	608	15	US-10-369-493-5748	Sequence 5748, Ap
33	573.5	17.0	695	12	US-10-424-599-176182	Sequence 176182,
34	571.5	17.0	623	12	US-10-424-599-154459	Sequence 154459,
35	570	16.9	559	15	US-10-369-493-5740	Sequence 5740, Ap
36	569.5	16.9	627	14	US-10-090-455-8	Sequence 8, Appli
37	569	16.9	604	9	US-09-745-763-197	Sequence 197, Appli
38	562.5	16.7	646	13	US-10-154-452-4	Sequence 4, Appli
39	561.5	16.7	819	12	US-10-425-114-54421	Sequence 54421, A
40	558.5	16.6	646	14	US-10-079-087-2	Sequence 2, Appli
41	555.5	16.5	646	14	US-10-090-455-13	Sequence 13, Appli
42	554.5	16.5	646	13	US-10-154-452-8	Sequence 8, Appli
43	554	16.4	610	15	US-10-369-493-5687	Sequence 5687, Ap
44	537.5	16.0	556	12	US-10-425-114-53846	Sequence 53846, A
45	537.5	16.0	673	12	US-10-425-114-64380	Sequence 64380, A

ALIGNMENTS

RESULT 1
US-09-837-992-1
; Sequence 1, Application US/09837992
; Patent No. US20020081687A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Hui
; APPLICANT: Schultz, Joshua
; APPLICANT: Shan, Bei
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: Sitolsterolemia Susceptibility Gene (SSG): Compositions and Methods of Use
; FILE REFERENCE: 018781-006020US
; CURRENT APPLICATION NUMBER: US/09/837,992
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 60/198,465
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 60/204,234
; PRIOR FILING DATE: 2000-05-15
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: mouse sitosterolemia susceptibility gene (SSG)
; OTHER INFORMATION: amino acid sequence
US-09-837-992-1

Query Match 100.0%; Score 3369; DB 9; Length 652;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 652; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGELPFLSPGARGPHINRGSLSLPQSGVTGTPEARHSLGLVHLVSVSNRVGFWNNIKS 60
Db 1 MGELPFLSPGARGPHINRGSLSLPQSGVTGTPEARHSLGLVHLVSVSNRVGFWNNIKS 60
Qy 61 CQOKWDRQILKDVSLYTESGQIMCILGSSGSGKTTLLDAISGLRLRTGTLEGEVFNCGE 120

Db 61 CQKWDQILKDVSLYIESQIMCIIIGSSGSGKTTLLDAISGRRLRTGTEGEVFNCGE 120
 Qy 121 LRQDFQDCFSYVQLQSDVFLSSLTRETLYATAMALCRSSADFYNNKVEAVMTLSLH 180
 Db 121 LRQDFQDCFSYVQLQSDVFLSSLTRETLYATAMALCRSSADFYNNKVEAVMTLSLH 180
 Qy 181 VADQMIGSYNFGGSSGERRRVSIAAQLQDPKVMMLDEPTTGLDCWTANQIVLLAELA 240
 Db 181 VADQMIGSYNFGGSSGERRRVSIAAQLQDPKVMMLDEPTTGLDCWTANQIVLLAELA 240
 Qy 241 RRDRIVITTHQPRSELFQHFEDKIALITYGELVFCGTPPEMLGFFNNCGYPCPEHSNPF 300
 Db 241 RRDRIVITTHQPRSELFQHFEDKIALITYGELVFCGTPPEMLGFFNNCGYPCPEHSNPF 300
 Qy 301 FYMDLTSDTQSREREIETTYKRVQMLECAFKESDIYHKILENIERARYLKTLPMPVFKTK 360
 Db 301 FYMDLTSDTQSREREIETTYKRVQMLECAFKESDIYHKILENIERARYLKTLPMPVFKTK 360
 Qy 361 DPPGMFGKGLVLLRRVTNLMRNKQAVIMRLVQNLINGLFLIFYLLRVQNNTLKGAVQDR 420
 Db 361 DPPGMFGKGLVLLRRVTNLMRNKQAVIMRLVQNLINGLFLIFYLLRVQNNTLKGAVQDR 420
 Qy 421 VGLLYQLVGATPYTGMLNAVNLFPMLRAVSDQSDGLYHKQWMLLAYVHLVLPFSVIAT 480
 Db 421 VGLLYQLVGATPYTGMLNAVNLFPMLRAVSDQSDGLYHKQWMLLAYVHLVLPFSVIAT 480
 Qy 481 VIFSSVCVWTGLYPEVARFGYFSAALLAPHLIGEFLLTVLLGIVQNPVNSIVALLSI 540
 Db 481 VIFSSVCVWTGLYPEVARFGYFSAALLAPHLIGEFLLTVLLGIVQNPVNSIVALLSI 540
 Qy 541 SGLLIGSGFIRNIQEMPIKILGYFTFKYCCCEILVNVNBYGLNFTCGSNTSMLNHPM 600
 Db 541 SGLLIGSGFIRNIQEMPIKILGYFTFKYCCCEILVNVNBYGLNFTCGSNTSMLNHPM 600
 Qy 601 CAITQGVQFIKTCGATSRFTANFLIYGIFIPALVILGIVIFKRDYLSR 652
 Db 601 CAITQGVQFIKTCGATSRFTANFLIYGIFIPALVILGIVIFKRDYLSR 652

RESULT 2
 US-09-989-981A-2
 ; Sequence 2, Application US/0989981A
 ; Publication No. US2003049730A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hobbs, Helen H.
 ; APPLICANT: Shan, Bei
 ; APPLICANT: Barnes, Robert
 ; APPLICANT: Tian, Hui
 ; APPLICANT: Tularik Inc.
 ; APPLICANT: Board of Regents, The University of Texas System
 ; TITLE OF INVENTION: ABCG5 and ABCG8: Compositions and Methods of Use
 ; FILE REFERENCE: 018781-007320US
 ; CURRENT APPLICATION NUMBER: US/09/989,981A
 ; CURRENT FILING DATE: 2002-07-23
 ; PRIOR APPLICATION NUMBER: US 60/252,235
 ; PRIOR FILING DATE: 2000-11-20
 ; PRIOR APPLICATION NUMBER: US 60/253,645
 ; PRIOR FILING DATE: 2000-11-26
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 652
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; FEATURE:
 ; OTHER INFORMATION: mouse ABCG5 (mABCG5)
 US-09-989-981A-2
 Query Match 100.0%; Score 3369; DB 10; Length 652;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 652; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGELPFLSPGARGPHINRGSLSEEQSVTGTEARHSLGVLHVSYSVSNRVGPFWMNIKS 60
 Db 1 MGELPFLSPGARGPHINRGSLSEEQSVTGTEARHSLGVLHVSYSVSNRVGPFWMNIKS 60
 Qy 61 CQKWDQILKDVSLYIESQIMCIIIGSSGSGKTTLLDAISGRRLRTGTEGEVFNCGE 120
 Db 61 CQKWDQILKDVSLYIESQIMCIIIGSSGSGKTTLLDAISGRRLRTGTEGEVFNCGE 120
 Qy 121 LRQDFQDCFSYVQLQSDVFLSSLTRETLYATAMALCRSSADFYNNKVEAVMTLSLH 180
 Db 121 LRQDFQDCFSYVQLQSDVFLSSLTRETLYATAMALCRSSADFYNNKVEAVMTLSLH 180
 Qy 181 VADQMIGSYNFGGSSGERRRVSIAAQLQDPKVMMLDEPTTGLDCWTANQIVLLAELA 240
 Db 181 VADQMIGSYNFGGSSGERRRVSIAAQLQDPKVMMLDEPTTGLDCWTANQIVLLAELA 240
 Qy 241 RRDRIVITTHQPRSELFQHFEDKIALITYGELVFCGTPPEMLGFFNNCGYPCPEHSNPF 300
 Db 241 RRDRIVITTHQPRSELFQHFEDKIALITYGELVFCGTPPEMLGFFNNCGYPCPEHSNPF 300
 Qy 301 FYMDLTSDTQSREREIETTYKRVQMLECAFKESDIYHKILENIERARYLKTLPMPVFKTK 360
 Db 301 FYMDLTSDTQSREREIETTYKRVQMLECAFKESDIYHKILENIERARYLKTLPMPVFKTK 360
 Qy 361 DPPGMFGKGLVLLRRVTNLMRNKQAVIMRLVQNLINGLFLIFYLLRVQNNTLKGAVQDR 420
 Db 361 DPPGMFGKGLVLLRRVTNLMRNKQAVIMRLVQNLINGLFLIFYLLRVQNNTLKGAVQDR 420
 Qy 421 VGLLYQLVGATPYTGMLNAVNLFPMLRAVSDQSDGLYHKQWMLLAYVHLVLPFSVIAT 480
 Db 421 VGLLYQLVGATPYTGMLNAVNLFPMLRAVSDQSDGLYHKQWMLLAYVHLVLPFSVIAT 480
 Qy 481 VIFSSVCVWTGLYPEVARFGYFSAALLAPHLIGEFLLTVLLGIVQNPVNSIVALLSI 540
 Db 481 VIFSSVCVWTGLYPEVARFGYFSAALLAPHLIGEFLLTVLLGIVQNPVNSIVALLSI 540
 Qy 541 SGLLIGSGFIRNIQEMPIKILGYFTFKYCCCEILVNVNBYGLNFTCGSNTSMLNHPM 600
 Db 541 SGLLIGSGFIRNIQEMPIKILGYFTFKYCCCEILVNVNBYGLNFTCGSNTSMLNHPM 600
 Qy 601 CAITQGVQFIKTCGATSRFTANFLIYGIFIPALVILGIVIFKRDYLSR 652
 Db 601 CAITQGVQFIKTCGATSRFTANFLIYGIFIPALVILGIVIFKRDYLSR 652

RESULT 3
 US-09-837-992-3
 ; Sequence 3, Application US/09837992
 ; Patent No. US20020081687A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tian, Hui
 ; APPLICANT: Schultz, Joshua
 ; APPLICANT: Shan, Bei
 ; APPLICANT: Tularik Inc.
 ; TITLE OF INVENTION: Sickle cell anemia Susceptibility Gene (SSG): Compositions
 ; FILE REFERENCE: 018781-006020US
 ; CURRENT APPLICATION NUMBER: US/09/837,992
 ; CURRENT FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: US 60/198,465
 ; PRIOR FILING DATE: 2000-04-18
 ; PRIOR APPLICATION NUMBER: US 60/204,234
 ; PRIOR FILING DATE: 2000-05-15
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 651
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: human sickle cell anemia susceptibility gene (SSG)
 ; OTHER INFORMATION: amino acid sequence
 US-09-837-992-3

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human ABCG5 (hABCG5)
US-09-989-981A-6

Query Match      81.5%; Score 2744.5; DB 9; Length 651;
Best Local Similarity 80.2%; Pred. No. 1.8e-258;
Matches 523; Conservative 64; Mismatches 64; Indels 1; Gaps 1;

QY 1 MGELPFLSPGARGPHINRGSLSLEQSGVTGTBARHSLGLVHVSYSVSNRVGPMWNKIS 60
DB 1 MGDLSLTPGSMGLQVNRGSSQSLGAPATAPEP-HSLGILHASYSVSHRVRPMDITS 59
QY 61 CQKWDRQILKDVSLYIESGQIMCILEQSGVTGTBARHSLGLVHVSYSVSNRVGPMWNKIS 120
DB 60 CRQWTRQILKDVSLYIESGQIMCILEQSGVTGTBARHSLGLVHVSYSVSHRVRPMDITS 119
QY 121 LRRDQFQDCFSYVLOSDFVLSLTVRETLRYTAMALCRSSADFYNNKVEAVMTLSLH 180
DB 120 LRRDQFQDCFSYVLOSDFVLSLTVRETLRYTAMALCRSSADFYNNKVEAVMTLSLH 179
QY 181 VADQMIGSYNFGGSSGERRRVSIAAQLQDPKVMLEDEPTTGLDCMTANQIVLLAELA 240
DB 180 VADRLIGNYSLGGISGERRRVSIAAQLQDPKVMLEDEPTTGLDCMTANQIVLLAELA 239
QY 241 RRDRIVIVTIHOPRSELFOHFKIAITYGELVFCGTPEEMLGFNNCGYPCPEHSNPF 300
DB 240 RNRRIVIVTIHOPRSELFOHFKIAITYGELVFCGTPEEMLGFNNCGYPCPEHSNPF 299
QY 301 FYMDLTSVDTQSKEREIETSKRVQIESAYKSAICHKTLKNIERMKHLKTLPMVPFKTK 360
DB 300 FYMDLTSVDTQSKEREIETSKRVQIESAYKSAICHKTLKNIERMKHLKTLPMVPFKTK 359
QY 361 DPPGMFGKLVLLRRVTRNLMRNKQAVIMRLVQNLIMGLFIYLLRVQNNTLKGAQDR 420
DB 360 DSPGVFSKLVLLRRVTRNLMRNKQAVIMRLVQNLIMGLFIYLLRVQNNTLKGAQDR 419
QY 421 VGLLYQLVGPATPYTGMLNANVLPMLRAVSDQESODGLYHKQWMLLAYVHLVPSVIAT 480
DB 420 VGLLYQLVGPATPYTGMLNANVLPMLRAVSDQESODGLYHKQWMLLAYVHLVPSVIAT 479
QY 481 VFSSVCYWTGLYHPEVARFGYFSAALLAPHILIGFELTLVLLGIVQNPVNSVALLSI 540
DB 480 MIFSSVCYWTGLYHPEVARFGYFSAALLAPHILIGFELTLVLLGIVQNPVNSVALLSI 539
QY 541 SGLLIGSGFIRNIQEMPIPKILGYFTFOKYCCCEILVWNEFYGLNFTCGGNTSMLNHPM 600
DB 540 AGVLVSGFLRNQIEMPIPKIISYFTFOKYCCCEILVWNEFYGLNFTCGGNTSMLNHPM 599
QY 601 CAITQGVQFIKTCPGATSRFTANFLIYGFIPALVILGIVIPKVRDYLIISR 652
DB 600 CAFTQGIQFIKTCPGATSRFTANFLIYGFIPALVILGIVIPKVRDYLIISR 651

```

```

RESULT 4
US-09-989-981A-6
; Sequence 6, Application US/099898981A
; Publication No. US20030049730A1
; GENERAL INFORMATION:
; APPLICANT: Hobbs, Helen H.
; APPLICANT: Shan, Bei
; APPLICANT: Barnes, Robert
; APPLICANT: Tian, Hui
; APPLICANT: Tularik Inc.
; APPLICANT: Board of Regents, The University of Texas System
; TITLE OF INVENTION: ABCG5 and ABCG8: Compositions and Methods of Use
; FILE REFERENCE: 018781-007320US
; CURRENT APPLICATION NUMBER: US/09/989,981A
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/252,235
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/253,645
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 651

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human ABCG5 (hABCG5)
US-09-989-981A-6

Query Match      81.5%; Score 2744.5; DB 10; Length 651;
Best Local Similarity 80.2%; Pred. No. 1.8e-258;
Matches 523; Conservative 64; Mismatches 64; Indels 1; Gaps 1;

QY 1 MGELPFLSPGARGPHINRGSLSLEQSGVTGTBARHSLGLVHVSYSVSNRVGPMWNKIS 60
DB 1 MGDLSLTPGSMGLQVNRGSSQSLGAPATAPEP-HSLGILHASYSVSHRVRPMDITS 59
QY 61 CQKWDRQILKDVSLYIESGQIMCILEQSGVTGTBARHSLGLVHVSYSVSNRVGPMWNKIS 120
DB 60 CRQWTRQILKDVSLYIESGQIMCILEQSGVTGTBARHSLGLVHVSYSVSHRVRPMDITS 119
QY 121 LRRDQFQDCFSYVLOSDFVLSLTVRETLRYTAMALCRSSADFYNNKVEAVMTLSLH 180
DB 120 LRRDQFQDCFSYVLOSDFVLSLTVRETLRYTAMALCRSSADFYNNKVEAVMTLSLH 179
QY 181 VADQMIGSYNFGGSSGERRRVSIAAQLQDPKVMLEDEPTTGLDCMTANQIVLLAELA 240
DB 180 VADRLIGNYSLGGISGERRRVSIAAQLQDPKVMLEDEPTTGLDCMTANQIVLLAELA 239
QY 241 RRDRIVIVTIHOPRSELFOHFKIAITYGELVFCGTPEEMLGFNNCGYPCPEHSNPF 300
DB 240 RNRRIVIVTIHOPRSELFOHFKIAITYGELVFCGTPEEMLGFNNCGYPCPEHSNPF 299
QY 301 FYMDLTSVDTQSKEREIETSKRVQIESAYKSAICHKTLKNIERMKHLKTLPMVPFKTK 360
DB 300 FYMDLTSVDTQSKEREIETSKRVQIESAYKSAICHKTLKNIERMKHLKTLPMVPFKTK 359
QY 361 DPPGMFGKLVLLRRVTRNLMRNKQAVIMRLVQNLIMGLFIYLLRVQNNTLKGAQDR 420
DB 360 DSPGVFSKLVLLRRVTRNLMRNKQAVIMRLVQNLIMGLFIYLLRVQNNTLKGAQDR 419
QY 421 VGLLYQLVGPATPYTGMLNANVLPMLRAVSDQESODGLYHKQWMLLAYVHLVPSVIAT 480
DB 420 VGLLYQLVGPATPYTGMLNANVLPMLRAVSDQESODGLYHKQWMLLAYVHLVPSVIAT 479
QY 481 VFSSVCYWTGLYHPEVARFGYFSAALLAPHILIGFELTLVLLGIVQNPVNSVALLSI 540
DB 480 MIFSSVCYWTGLYHPEVARFGYFSAALLAPHILIGFELTLVLLGIVQNPVNSVALLSI 539
QY 541 SGLLIGSGFIRNIQEMPIPKILGYFTFOKYCCCEILVWNEFYGLNFTCGGNTSMLNHPM 600
DB 540 AGVLVSGFLRNQIEMPIPKIISYFTFOKYCCCEILVWNEFYGLNFTCGGNTSMLNHPM 599
QY 601 CAITQGVQFIKTCPGATSRFTANFLIYGFIPALVILGIVIPKVRDYLIISR 652
DB 600 CAFTQGIQFIKTCPGATSRFTANFLIYGFIPALVILGIVIPKVRDYLIISR 651

```

```

RESULT 5
US-10-090-455-6
; Sequence 6, Application US/10090455
; Publication No. US20030027259A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hongyun
; APPLICANT: Le Bihan, Stephane
; TITLE OF INVENTION: NOVEL ABCG4 TRANSPORTER AND USES THEREOF
; FILE REFERENCE: 100103.406
; CURRENT APPLICATION NUMBER: US/10/090,455
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-090-455-6

```

Query Match 81.5%; Score 2744.5; DB 14; Length 651;
 Best Local Similarity 80.2%; Pred. No. 1.8e-258;
 Matches 523; Conservative 64; Mismatches 64; Indels 1; Gaps 1;

QY 1 MGELPFLSPGARGPHINRGSLSLGSGSVTGTAEARHSLGVLHVSVSNRVCPPWNKIS 60
 DB 1 MGDLSUTPGSGMLQVNRGSLGSLGAGAPATAPEP-HSLGILHASVSVSRVPMWDITS 59

QY 61 CQQRWDRLKDVLYIESGGIMCILGSSGSGTKTLLDAISGRRLRTGTLEGEVFNCGE 120
 DB 60 CQQRWDRLKDVLYIESGGIMCILGSSGSGTKTLLDAMSGRLGRAGTLEGEVYNGRA 119

QY 121 LRRQFQDCRSYVLOSDFVLSLTVRETLAYTAMALCRSSADFYNNKKEAVMTLSLH 180
 DB 120 LRRQFQDCRSYVLOSDFVLSLTVRETLAYTAMALCRSSADFYNNKKEAVMTLSLH 179

QY 181 VADQMIGSYNFGGSSGERRRVSIAAQLQDPKVMMLDEPTTGLDCMTANQIVLLAELA 240
 DB 180 VADRLIGNYSIGGISTGERRRVSIAAQLQDPKVMMLDEPTTGLDCMTANQIVLLAELA 239

QY 241 RRDRIIVITHQPRSELFQHPDKIAITYGELVFCGTPPEMLGFFNNCGYPCPEHSNPF 300
 DB 240 RNRIRIVLTHQPRSELFQHPDKIAITYGELVFCGTPPEMLGFFNNCGYPCPEHSNPF 299

QY 301 FYMDLTSVDTOSRRETIETKRVQMLECAPKESDIYHKILENIERARYLXTLPMVPPKTK 360
 DB 300 FYMDLTSVDTOSRRETIETKRVQMLECAPKESDIYHKILENIERARYLXTLPMVPPKTK 359

QY 361 DPPQMGKGLVLLRRVTRNLMNRQAVIMRLVQNLINGLFLIFVLLRVQNTLKGAVQDR 420
 DB 360 DSPGVFSKGLVLLRRVTRNLMNRQAVIMRLVQNLINGLFLIFVLLRVQNTLKGAVQDR 419

QY 421 VGLLYQLVGATPYTGMNLAVNLPMLRAVSDQSDGLYHKWQMLLAYVHLVLPFSVIAT 480
 DB 420 VGLLYQLVGATPYTGMNLAVNLPMLRAVSDQSDGLYHKWQMLLAYVHLVLPFSVIAT 479

QY 481 VIFSSVCYMTLGLYPEVARFGYFSAALLAPHLIGEFLLTVLLGVQNPVNIVSIVALLSI 540
 DB 480 MIFSSVCYMTLGLYPEVARFGYFSAALLAPHLIGEFLLTVLLGVQNPVNIVSIVALLSI 539

QY 541 SGLLIGSFIRNIOEMPIPLKILGVTFQKCCILVNVFYNFTCGGNTSMNLNHPM 600
 DB 540 AGVLVSGFLRNIOEMPIPLKILGVTFQKCCILVNVFYNFTCGGNTSMNLNHPM 599

QY 601 CAITQGVQFIKTCFGATSRFTANFLIYGFIPALVILGIVIFKVRDYLSR 652
 DB 600 CAITQGVQFIKTCFGATSRFTANFLIYGFIPALVILGIVIFKVRDYLSR 651

RESULT 6
 US-10-104-047-2795
 ; Sequence 2795, Application US/10104047
 ; Publication No. US20030236392A1
 ; GENERAL INFORMATION:
 ; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
 ; FILE REFERENCE: H1-A0105
 ; CURRENT APPLICATION NUMBER: US/10/104,047
 ; CURRENT FILING DATE: 2002-03-25
 ; PRIOR APPLICATION NUMBER:
 ; PRIOR FILING DATE:
 ; NUMBER OF SEQ ID NOS: 4096
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2795
 ; LENGTH: 256
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-104-047-2795

Query Match 34.9%; Score 1177; DB 15; Length 256;
 Best Local Similarity 85.5%; Pred. No. 4.6e-106;
 Matches 219; Conservative 23; Mismatches 14; Indels 0; Gaps 0;

QY 397 MGLFLIFVLLRVQNTLKGAVQDRVGLLYQLVGATPYTGMNLAVNLPMLRAVSDQSD 456
 DB 1 MGLFLIFVLLRVQNTLKGAVQDRVGLLYQLVGATPYTGMNLAVNLPMLRAVSDQSD 60

QY 457 GLXHKWQMLLAYVHLVLPFSVIATVIFSSVCYMTLGLYPEVARFGYFSAALLAPHLIGEF 516
 DB 61 GLYQKQWMLLAYVHLVLPFSVIATVIFSSVCYMTLGLYPEVARFGYFSAALLAPHLIGEF 120

QY 517 LTVLLGVQNPVNIVSIVALLISGLLIGSGFTENQEMPIPLKILGYFTFOKVCCEIL 576
 DB 121 LTVLLGVQNPVNIVSIVALLISGLLIGSGFTENQEMPIPLKILGYFTFOKVCCEIL 180

QY 577 VVNEFYGLNFTCGGNTSMNLNHPMCAITQGVQFIKTCFGATSRFTANFLIYGFIPALV 636
 DB 181 VVNEFYGLNFTCGGNTSMNLNHPMCAITQGVQFIKTCFGATSRFTANFLIYGFIPALV 240

QY 637 ILGIVIFKVRDYLSR 652
 DB 241 ILGIVIFKVRDYLSR 256

RESULT 7
 US-09-989-981A-4
 ; Sequence 4, Application US/09989981A
 ; Publication No. US20030049730A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hobbs, Helen H.
 ; APPLICANT: Shan, Bei
 ; APPLICANT: Barnes, Robert
 ; APPLICANT: Tian, Hui
 ; APPLICANT: Tullarik Inc.
 ; TITLE OF INVENTION: ABCG8 and ABCG8: Compositions and Methods of Use
 ; FILE REFERENCE: 018781-007320US
 ; CURRENT APPLICATION NUMBER: US/09/989,981A
 ; CURRENT FILING DATE: 2002-07-23
 ; PRIOR APPLICATION NUMBER: US 60/252,235
 ; PRIOR FILING DATE: 2000-11-20
 ; PRIOR APPLICATION NUMBER: US 60/253,645
 ; PRIOR FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 672
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; FEATURE:
 ; OTHER INFORMATION: mouse ABCG8 (mABCG8)
 US-09-989-981A-4

Query Match 20.8%; Score 701.5; DB 10; Length 672;
 Best Local Similarity 29.1%; Pred. No. 5.5e-59;
 Matches 194; Conservative 131; Mismatches 245; Indels 97; Gaps 19;

QY 27 QGSVTGTAEARHSLGVLHVSVS-----VSNRVGPW-----WNTKS 60
 DB 24 QDSLFSSESDNS--LYFTYSGQSNLTLEVRDLTYQVDIASQV-PWFEQLAQPKIPWRS 79

QY 61 CQQRWDRLKDVLYIESGGIMCILGSSGSGTKTLLDAISGRRLRTGTLEGEVFNCGE 120
 DB 80 SQDSCELGI-RNLSFKVRSGQMLAIGSSGCGRSLLDVITGRHGKMKSCQIWIINGQP 138

QY 121 LRRDQFQDCRSYVLOSDFVLSLTVRETLAYTAMALCRS-SADFYNKKEAVMTLSLS 179
 DB 139 STPOLVRKCVAVRHQDQLLPNLTVRETLAFIAQMLPRTFSQAQDKRVEDIAELRLR 198

QY 180 HVADQMIGSYNFGGSSGERRRVSIAAQLQDPKVMMLDEPTTGLDCMTANQIVLLAEL 239
 DB 199 QCANTRVGNVTYVRGSGGERRRVSIGVQLLWNPGLILDEPTSGLDSTFAHNLVTLSRL 258

QY 240 ARDRIVIVITHQPRSELFQHPDKIAITYGELVFCGTPPEMLGFFNNCGYPCPEHSNPF 299
 DB 241 ARDRIVIVITHQPRSELFQHPDKIAITYGELVFCGTPPEMLGFFNNCGYPCPEHSNPF 299

Db 259 AKGNRLVLSLHQPASDIFRLFDVLVLTSGTPIYLGAQOMVQYFSTIGHPCPYSNPA 318
 Qy 300 DFYMDLTSVDTQSREREIETRYQVOMLECAFKP-----SDIYHKI-LNTERARYLTKLP 353
 Db 319 DFYVDTLSIDRSKEREVATVEKAQSLAALFLEKVQGFDDFLWKABEAKELNTSHTVSLT 378
 Qy 354 MVPFKTQDP-----PGMEGKLVLLRRVTRNLNENKQAVINRLVQNLIMGLFLIPLYL 405
 Db 379 L-----TQDTCGTAVELPGMTIEQSTLIRROISNDFRDLPTLLIHGSEACMSLIIGP-- 432
 Qy 406 LRQVNTLKGAVQDRVGLLYQVATPYTGMNLAVNLPMLRAVSDOESODGLYHKWQML 465
 Db 433 LVYHGAKQLSFMDDTAALLFMIGALIPNVILDVVSKCHSRSMLYVELEDGLYTAGPYF 492
 Qy 466 LAYVLHVLFPFVIATVIFSSVCYWTILGLYPEVAPGYFSAALLAPHLIGFEL-----TL 519
 Db 493 FAKILGELPEHCAYVYIYAMPYITNLRPVPELF-----LL-HELLVWLWVFCRTM 544
 Qy 520 VLLGIVQNPNT-VNSIVALLSIGLLIGSGFIRNIQEMPIPKILGYFTTQKYCEILVV 578
 Db 545 ALAASAMLPFTFMSSFFCNALYNSFYLTAGFMNLNLIWIPAMISKLSFLRWCFSLMQ 604
 Qy 579 NEFYGL-----NFTCGSNTSMJ-----NHPMCA---ITQGVQFIEKTCPGATSRFT 622
 Db 605 IQFNHGLYTTQIGNFTFSILGDTMISAMDLSNHPYIYILVIGISY-----651
 Qy 623 ANFLILY 629
 Db 652 -GPLEFLY 657

RESULT 8
 US-09-961-086-1
 ; Sequence 1, Application US/09961086
 ; Publication No. US20030036645A1
 ; GENERAL INFORMATION:
 ; APPLICANT: UNIVERSITY OF MARYLAND, BALTIMORE
 ; APPLICANT: ROSS, Douglas D.
 ; APPLICANT: DOYLE, L. Austin
 ; APPLICANT: ABRUZZO, Lynne
 ; TITLE OF INVENTION: BREAST CANCER RESISTANCE PROTEIN (BCRP) AND THE DNA
 ; TITLE OF INVENTION: WHICH ENCODES IT
 ; FILE REFERENCE: EP19376-019
 ; CURRENT APPLICATION NUMBER: US/09/961,086
 ; CURRENT FILING DATE: 2001-09-21
 ; PRIOR APPLICATION NUMBER: US 60/073,763
 ; PRIOR FILING DATE: 1998-02-05
 ; PRIOR APPLICATION NUMBER: PCT/US99/02577
 ; PRIOR FILING DATE: 1999-02-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 655
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-961-086-1

Query Match 20.6%; Score 693.5; DB 10; Length 655;
 Best Local Similarity 29.0%; Pred. No. 3.2e-58;
 Matches 181; Conservative 142; Mismatches 246; Indels 55; Gaps 16;
 Qy 25 LEQGSVTGTEARHS-----LGVLHVSVSNRVGPMWNIKSQQKWDRLQKDV 73
 Db 12 VSQNTNGPPATASNDLKAFTEGAVLSFNICYRVKLKSG----FLPCRKPVEKELLSNI 67
 Qy 74 SLVIESGQIMCIIIGSSGSKTLLDAISGLRRTGTLEGEVFNVCCELRRQFQDCFSYV 133
 Db 68 NGIMKPG-LNALIGPTGGKSLDLVLAARKDPG--LSGDVLINGAP-RPANFKCNSGYV 124
 Qy 134 LOSDVFLSLTRETIRYFAMALCRSSADF-YNKKVEAVMTLSLHSHVADQMIGSYNFG 192
 Db 125 VQDDVVMGTLTRENLPFAALRLATTMTNHEKNERINRVICQELGLDKVADSKVGTQFIR 184

Qy 193 GISSERRVSTAAQLQDPEKVMMLDEPTTGLDCTMTANQIVLLLAELARRDRIVVTIHO 252
 Db 185 GVSGERKRTSLGMBLITDPSILFDEFTTGLDSSANAVLLLLKRMKQGETIIFSIO 244
 Qy 253 PRSELFQHFDKIAIITYGELVFCGTPEEMGLFFNNCGYCPPEHSHNPFDFYMDLTSVDTQ- 311
 Db 245 PRYSIFKLPDLSLTLASGRMLPHGPAQEGALGYFESAGYHCEAYNNPADFFLDIINGDSTA 304
 Qy 312 -SREREIETRYQVOMLECAFKPSDIYHKI-----LNTERARYLTK 351
 Db 305 VALNRE-EDFKATEIIEPSKQDKPLIEKLAETVNSSYKTKABELHQSOGGKKKKTIV 363
 Qy 352 LPMVPFKTKDPPGMFGKLGVLRLVTRNLNENKQAVINRLVQNLIMGLFLIPLYLVRQNN 411
 Db 364 PREISYTT-----SFCHQLRWVSKRSFKNLGNPQASIAQLIIVTVVLGLVIGAIYFGLKND 419
 Qy 412 TKGAVQDRVGLLYQVATPYTGMNLAVNLPMLRAVSDOESODGLYHKWQMLLAYVL- 470
 Db 420 ST-GIGNAGVLFLLTNNQCFSS-VSAVELFVVEKKLFIHEYISGYRVSSYFGLKLLS 476
 Qy 471 HVLPPSVIATVIFSSVCYWTILGLYPEVAPGYFSAALLAPHLIGFELTLVLLGIVQNPNI 530
 Db 477 DLLPMTMLPSLIIFTCTIVFMLGLKPKADAFVMMFTLM--MVAYSASSMALAIAAGQSV 533
 Qy 531 VNSIVALLSIS--GLLIGSGFIRNIQEMPIPKILGYFTTQKYCEILVNVNPFYGLNFTC 588
 Db 534 VSVATLLMTICFVFMWIFSGLLVNLTTIASWLSWLYQFSPRYGFTALQHNFLQNF-C 592
 Qy 589 GGSNTSMNLNHPMCAITQGVQFIEK 612
 Db 593 PGLNATGNPNCAVCTGBEYLVK 616

RESULT 9
 US-10-405-806-13
 ; Sequence 13, Application US/10405806
 ; Publication No. US2003023262A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KONATANI, HIDEYA
 ; APPLICANT: HARA, YOSHIKAZU
 ; APPLICANT: KOTANI, HIDEHITO
 ; APPLICANT: NAKAGAWA, RINAKO
 ; TITLE OF INVENTION: DRUG RESISTANT GENE AND USE THEREOF
 ; FILE REFERENCE: 234985USOCONT
 ; CURRENT APPLICATION NUMBER: US/10/405,806
 ; CURRENT FILING DATE: 2003-04-03
 ; PRIOR APPLICATION NUMBER: PCT/JP01/08112
 ; PRIOR FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: JP2000-303441
 ; PRIOR FILING DATE: 2000-10-03
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 13
 ; LENGTH: 655
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: ABCG2 482Tmutant sequence
 US-10-405-806-13

Query Match 20.6%; Score 693.5; DB 15; Length 655;
 Best Local Similarity 29.0%; Pred. No. 3.2e-58;
 Matches 181; Conservative 142; Mismatches 246; Indels 55; Gaps 16;
 Qy 25 LEQGSVTGTEARHS-----LGVLHVSVSNRVGPMWNIKSQQKWDRLQKDV 73
 Db 12 VSQNTNGPPATASNDLKAFTEGAVLSFNICYRVKLKSG----FLPCRKPVEKELLSNI 67
 Qy 74 SLVIESGQIMCIIIGSSGSKTLLDAISGLRRTGTLEGEVFNVCCELRRDQDCFSYV 133
 Db 68 NGIMKPG-LNALIGPTGGKSLDLVLAARKDPG--LSGDVLINGAP-RPANFKCNSGYV 124
 Qy 134 LOSDVFLSLTRETIRYFAMALCRSSADF-YNKKVEAVMTLSLHSHVADQMIGSYNFG 192

```

125 VQDDVVGTTVRENLOFSAAALRLATMTNHEKNERINRVIQELGDLKVDKSVGTQFIR 184
193 GISGERRRVSIAAQLLODPKVMMLDEPTTGLDQMTANQIVLLAELARRDRIVIVTIHQ 252
185 GVSGERKRTSIGMELITDPSILFDPTTGLDSTANAVLLLLKMSKQGRITIFSIHQ 244
253 PRSELFQHFDAKIALTYGELVFCGTPEMLGFFNCGYPCPEHSNPPDFYMDLTSVDTO- 311
245 PRYSIFKLFDSLTLLASGRMLFHGPAQCALGYFESAGYHCEAYNNPADFFLDIINGDSTA 304
312 -SRREIETKYRVQMLECAPKESDIYHKI-----LENIERARYLKT 351
305 VALNRE-EDFKATEIIBPSKQDKPLIEKLABIYVNSSFYKETAELHQLSGGKKKITV 363
352 LPMVFFTKDPPGMFGKGLVLLRRVTRNLMNKQAVIMLVQNLIMGLFLIFYLLRVQNN 411
364 FKEISYTT----SFCHQLRWVSKSFKNLLGNPOASIAQIIVTVLGLVIGALYFGLKND 419
412 TLKGAQVDRVGLLYQLVGATPYTGMNAVNLFPMRAVSDQSDGLYHKWQMLLAYVL- 470
420 ST--GIQNRAGVLPFLTTNQCFSS-VSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLS 476
471 HVLPSVIATVIFSSVCYWTILGLYPEVARFGYFSAALLAPHLIGEFTLVLLGIVQNPNI 530
477 DLLPMTLPSIIFTCIVYFVLMGLKPKADAFVMMFTLM--WVAYSASSMALAIAAGQSV 533
531 VNSIVALLSIS--GLLTGSGFIRNIQEMPIPLKILGYFTFKYCCCEILVNVFVGLNFTC 588
534 VSVATLLMTICFVFMWIFSGLLVNLTIASWLSWLYQFSPRYGFTALQHNEFLGQNF-C 592
589 GGSNTSMNLNHPMCAITQGVQPIEK 612
593 PGLNATGNPCNYATCTGEEYLVK 616

```

RESULT 10

```

US-09-981-353-35
; Sequence 35, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981.353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 35
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID NO. US20020160382A1 5517972CD1
US-09-981-353-35

```

```

Query Match 20.5%; Score 691.5; DB 9; Length 655;
Best Local Similarity 29.0%; Pred. No. 5e-58;
Matches 181; Conservative 141; Mismatches 247; Indels 55; Gaps 16;

QY 25 LEQGSVTGTAEARHS-----LGVLHVSYSVSNRVGFWPWNKSCQCKWDRQILKDV 73
DB 12 VSQNGTNGFPATASNDLKAFTEGAVLSFHNICRYVKLSG---FLPCRKPEVEKILSNI 67
QY 74 SLVIESGQIMCITLSSGSGKTTLLDAISGRLLRTGTLEGEVFNVCGLRRDQFCFSYV 133
DB 68 NQIMKPG-LNALGTGTGSGSLLDVLAARKDPSG-LSGDVLIAGP-RPANFKNSGYV 124
QY 134 LQSDVFLSRLTYRETLRYTAMALCRSADF-YNKKVEAVMTLSLSHVADQMGISYNG 192
DB 125 VQDDVVGTTVRENLOFSAAALRLATMTNHEKNERINRVIQELGDLKVDKSVGTQFIR 184

```

```

QY 193 GISGERRRVSIAAQLLODPKVMMLDEPTTGLDQMTANQIVLLAELARRDRIVIVTIHQ 252
DB 195 GVSGERKRTSIGMELITDPSILFDPTTGLDSTANAVLLLLKMSKQGRITIFSIHQ 244
QY 253 PRSELFQHFDAKIALTYGELVFCGTPEMLGFFNCGYPCPEHSNPPDFYMDLTSVDTO- 311
DB 245 PRYSIFKLFDSLTLLASGRMLFHGPAQCALGYFESAGYHCEAYNNPADFFLDIINGDSTA 304
QY 312 -SRREIETKYRVQMLECAPKESDIYHKI-----LENIERARYLKT 351
DB 305 VALNRE-EDFKATEIIBPSKQDKPLIEKLABIYVNSSFYKETAELHQLSGGKKKITV 363
QY 352 LPMVFFTKDPPGMFGKGLVLLRRVTRNLMNKQAVIMLVQNLIMGLFLIFYLLRVQNN 411
DB 364 FKEISYTT----SFCHQLRWVSKSFKNLLGNPOASIAQIIVTVLGLVIGALYFGLKND 419
QY 412 TLKGAQVDRVGLLYQLVGATPYTGMNAVNLFPMRAVSDQSDGLYHKWQMLLAYVL- 470
DB 420 ST--GIQNRAGVLPFLTTNQCFSS-VSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLS 476
QY 471 HVLPSVIATVIFSSVCYWTILGLYPEVARFGYFSAALLAPHLIGEFTLVLLGIVQNPNI 530
DB 477 DLLPMTLPSIIFTCIVYFVLMGLKPKADAFVMMFTLM--WVAYSASSMALAIAAGQSV 533
QY 531 VNSIVALLSIS--GLLTGSGFIRNIQEMPIPLKILGYFTFKYCCCEILVNVFVGLNFTC 588
DB 534 VSVATLLMTICFVFMWIFSGLLVNLTIASWLSWLYQFSPRYGFTALQHNEFLGQNF-C 592
QY 589 GGSNTSMNLNHPMCAITQGVQPIEK 612
DB 593 PGLNATGNPCNYATCTGEEYLVK 616

```

RESULT 11

```

US-10-120-687-61
; Sequence 61, Application US/10120687
; Publication No. US20030082155A1
; GENERAL INFORMATION:
; APPLICANT: Massachusetts General Hospital
; TITLE OF INVENTION: Stem Cells of the Islets of Langerhans and Their Use in Treating
; FILE REFERENCE: 3284/1235B
; CURRENT APPLICATION NUMBER: US/10/120.687
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: US60/169082
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 09/963,875
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/215,109
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: US 60/238880
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 09/731261
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 61
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-120-687-61

```

```

Query Match 20.5%; Score 691.5; DB 14; Length 655;
Best Local Similarity 29.0%; Pred. No. 5e-58;
Matches 181; Conservative 141; Mismatches 247; Indels 55; Gaps 16;

QY 25 LEQGSVTGTAEARHS-----LGVLHVSYSVSNRVGFWPWNKSCQCKWDRQILKDV 73
DB 12 VSQNGTNGFPATASNDLKAFTEGAVLSFHNICRYVKLSG---FLPCRKPEVEKILSNI 67
QY 74 SLVIESGQIMCITLSSGSGKTTLLDAISGRLLRTGTLEGEVFNVCGLRRDQFCFSYV 133

```

Db 68 NGIMKPG-LNALIGPTGGKSSLLDLAARADPSG-LSGDVLINGAP-RPANFKCNSGV 124
 QY 134 LOSDFVLSLTVRETLRYTAMALCRSSADF-YNKKEAVMTLSLHVADQMISYNG 192
 Db 125 VODDVVMTLVRENLQFSAALRLATTWNEKNERINRVQELGDKVADSKVGTQFIR 184
 QY 193 GISSGERRVIAAQLQDPKVMMLDEPTTGLDCMTANOIVLLAEALARRDRIVVITHQ 252
 Db 185 GVSGERKRTSIMGELITDPSILFDDEPTTGLDSTANAVLLKRMKSKQRTIIFSIHQ 244
 QY 253 PRSELFQHDKIALITYGELVFCGTPPEMLGFENNGYPCPEHNSPDPFVMDLTSVDTQ- 311
 Db 245 PRYSIFKLFDSLTLASGRMLPHGPAQALGFESAGYHCEAYNNPADFFLDIINGDSTA 304
 QY 312 -SREIEITYKRVQMLECAFRESDIYHKI-----LENIERARYLKT 351
 Db 305 VALNRE-EDFKATEIIEPSKQDKPLIEKLAIEYVNSFYKETAELHQLSGGKKKITV 363
 QY 352 LPMVPKTKDPPGMFGKLVLLRRVTRNLMENKQAVIMRLVQNLIMGLFIYLLRVQNN 411
 Db 364 FKEISYTT----SFCQLRWVSKRSFKNLLGNPQASIAQIIVTVVLGIVGAIYFGLKND 419
 QY 412 TLKGAVQDRVGLLYQVLGATPYTGMNAVNLPMLRAVSDQESQDGLYHKQMMLLAYVL- 470
 Db 420 ST--GIQNRAGVLFELTTNQCPSS-VSAVELFVVEKKLFIEHYISGYRVSSYFLGKLS 476
 QY 471 HVLPPSVIATVPSSVCVWTGLYPEVARFGYFSAALLAPHLIGEBFLTLVLLGIVQNPNI 530
 Db 477 DLLPMRLPSIIFTCTIVFVLMGLKPKADAFVVMFTLM---MVAYSASSMALAIAAGQSV 533
 QY 531 VNSIVALLSIS--GLLIGSGFRNIQEMPIPLKILGYFTFOKYCCBILVNVNPFYGLNFTC 588
 Db 534 VSVATLLMTICFVFMIFSGLLVNLTTIASWLSWLOYSIPRYGFTALQHNEFLGQNF-C 592
 QY 589 GGSNTSMLNHPMCAITQGVQFIEK 612
 Db 593 PGLNATGNPCNYATCTGEEYLVK 616

RESULT 12

US-10-405-806-2
 ; Sequence 2, Application US/10405806
 ; Publication NO. US20030232362A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KOMATANI, HIDEVA
 ; APPLICANT: HARA, YOSHIKAZU
 ; APPLICANT: KOTANI, HIDEHITO
 ; APPLICANT: NAKAGAWA, KINAKO
 ; TITLE OF INVENTION: DRUG RESISTANT GENE AND USE THEREOF
 ; FILE REFERENCE: 234985USOCNT
 ; CURRENT APPLICATION NUMBER: US/10/405,806
 ; PRIOR FILING DATE: 2003-04-03
 ; PRIOR APPLICATION NUMBER: PCT/JP01/08112
 ; PRIOR FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: JP2000-303441
 ; PRIOR FILING DATE: 2000-10-03
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 2
 ; LENGTH: 655
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-405-806-2

Query Match 20.5%; Score 691.5; DB 15; Length 655;
 Best Local Similarity 29.0%; Pred. No. 5e-58;
 Matches 181; Conservative 141; Mismatches 247; Indels 55; Gaps 16;
 QY 25 LEQGSVTGTAEARHS-----LGVLHVSYSVSNRVGPPWNNIKSCQCKWDRQILKDV 73
 Db 12 VSQGNTNGFPATASNDLKAFTEGAVLSFPHNICYRVKLKSG-----FLPCRKPVEKILSNI 67
 QY 74 SLVIESQIMCILGSSGCKTLLDASGLRLRTGTLEGEVFNCGELRRDQFCFSV 133

Db 68 NGIMKPG-LNALIGPTGGKSSLLDLAARADPSG-LSGDVLINGAP-RPANFKCNSGV 124
 QY 134 LOSDFVLSLTVRETLRYTAMALCRSSADF-YNKKEAVMTLSLHVADQMISYNG 192
 Db 125 VODDVVMTLVRENLQFSAALRLATTWNEKNERINRVQELGDKVADSKVGTQFIR 184
 QY 193 GISSGERRVIAAQLQDPKVMMLDEPTTGLDCMTANOIVLLAEALARRDRIVVITHQ 252
 Db 185 GVSGERKRTSIMGELITDPSILFDDEPTTGLDSTANAVLLKRMKSKQRTIIFSIHQ 244
 QY 253 PRSELFQHDKIALITYGELVFCGTPPEMLGFENNGYPCPEHNSPDPFVMDLTSVDTQ- 311
 Db 245 PRYSIFKLFDSLTLASGRMLPHGPAQALGFESAGYHCEAYNNPADFFLDIINGDSTA 304
 QY 312 -SREIEITYKRVQMLECAFRESDIYHKI-----LENIERARYLKT 351
 Db 305 VALNRE-EDFKATEIIEPSKQDKPLIEKLAIEYVNSFYKETAELHQLSGGKKKITV 363
 QY 352 LPMVPKTKDPPGMFGKLVLLRRVTRNLMENKQAVIMRLVQNLIMGLFIYLLRVQNN 411
 Db 364 FKEISYTT----SFCQLRWVSKRSFKNLLGNPQASIAQIIVTVVLGIVGAIYFGLKND 419
 QY 412 TLKGAVQDRVGLLYQVLGATPYTGMNAVNLPMLRAVSDQESQDGLYHKQMMLLAYVL- 470
 Db 420 ST--GIQNRAGVLFELTTNQCPSS-VSAVELFVVEKKLFIEHYISGYRVSSYFLGKLS 476
 QY 471 HVLPPSVIATVPSSVCVWTGLYPEVARFGYFSAALLAPHLIGEBFLTLVLLGIVQNPNI 530
 Db 477 DLLPMRLPSIIFTCTIVFVLMGLKPKADAFVVMFTLM---MVAYSASSMALAIAAGQSV 533
 QY 531 VNSIVALLSIS--GLLIGSGFRNIQEMPIPLKILGYFTFOKYCCBILVNVNPFYGLNFTC 588
 Db 534 VSVATLLMTICFVFMIFSGLLVNLTTIASWLSWLOYSIPRYGFTALQHNEFLGQNF-C 592
 QY 589 GGSNTSMLNHPMCAITQGVQFIEK 612
 Db 593 PGLNATGNPCNYATCTGEEYLVK 616

RESULT 13

US-09-866-866A-10
 ; Sequence 10, Application US/09866866A
 ; Patent No. US20020102244A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sorrentino, Brian
 ; APPLICANT: Schuetz, John
 ; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
 ; FILE REFERENCE: 1340-1-021CIP2
 ; CURRENT APPLICATION NUMBER: US/09/866,866A
 ; CURRENT FILING DATE: 2001-08-30
 ; PRIOR FILING DATE: 2001-05-31
 ; PRIOR APPLICATION NUMBER: PCT/US99/11825
 ; PRIOR FILING DATE: 1999-05-27
 ; PRIOR APPLICATION NUMBER: 60/086,988
 ; PRIOR FILING DATE: 1998-05-28
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 10
 ; LENGTH: 655
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 US-09-866-866A-10

Query Match 20.5%; Score 689.5; DB 9; Length 655;
 Best Local Similarity 29.0%; Pred. No. 7.9e-58;
 Matches 181; Conservative 141; Mismatches 247; Indels 55; Gaps 16;
 QY 25 LEQGSVTGTAEARHS-----LGVLHVSYSVSNRVGPPWNNIKSCQCKWDRQILKDV 73
 Db 12 VSQGNTNGFPATVSNLDKAFTEGAVLSFPHNICYRVKLKSG-----FLPCRKPVEKILSNI 67

